Introduction to GIS

through Quantum GIS

Day 2: Technical Stuff Day

offered by Valley Stewardship Network

taught and prepared by Legion GIS, LLC

July 2016

Today's Technical Stuff…

Spatial Data

--vector vs. raster

--common formats of each category

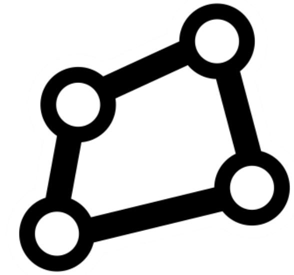
(remember, this whole GIS thing is just data storage and interpretation)

Coordinate Reference Systems

--needed to correctly interpret the spatial data

(spatial data without a coordinate reference system is as misunderstood as a moody adolescent)

vector data



- Stored as discrete vector geometries, called “features”
- Three basic types of geometries: point, line, polygon
  - There are also more complex variations like “multipoint”, or “multipolygon”
- An **attribute table** accompanies the geometries, with one row per feature
- The attribute table can have many columns (or “fields”), each containing a piece of information about the feature
- In QGIS, use the Identify tool to quickly view all attributes for a given feature

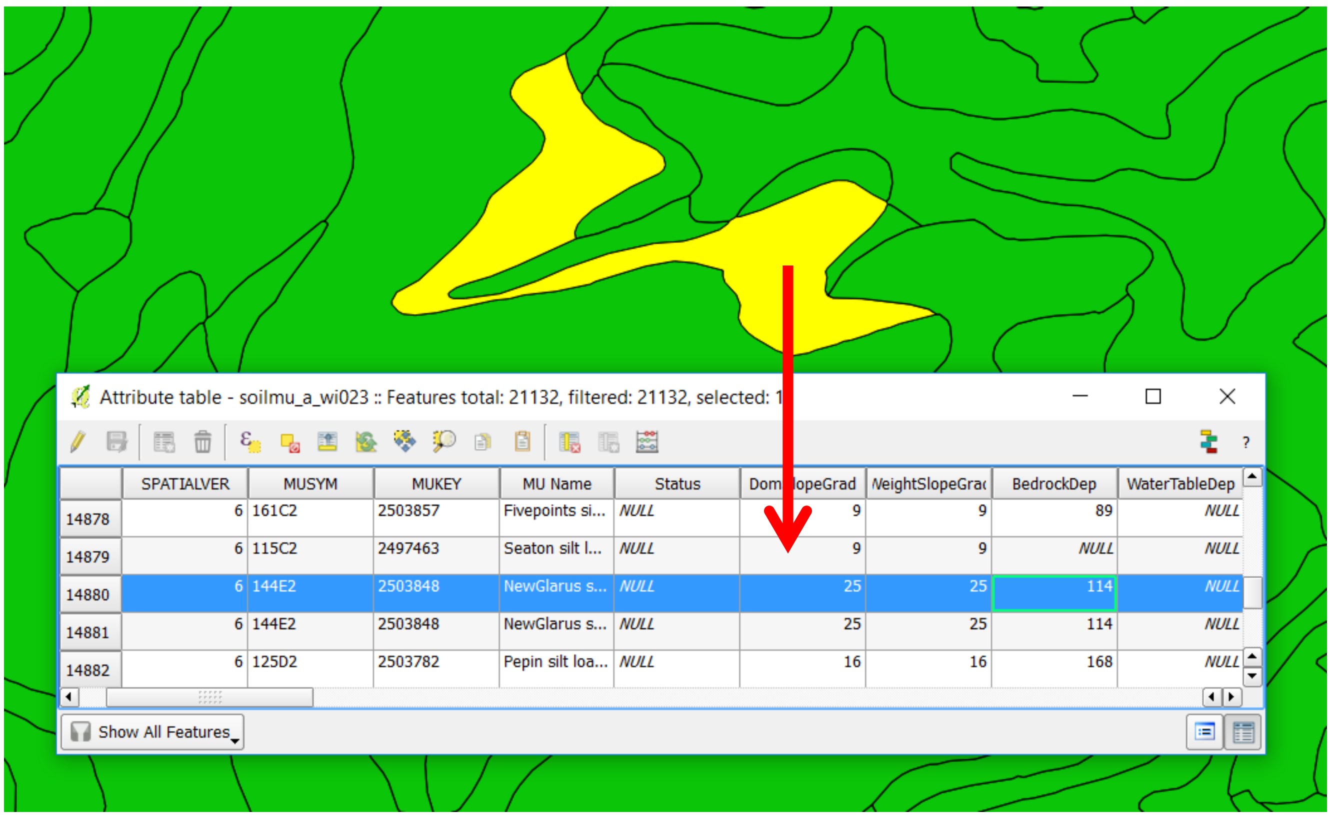
example of vector data…

the geometry

(these are soil unit polygons)

the attribute table

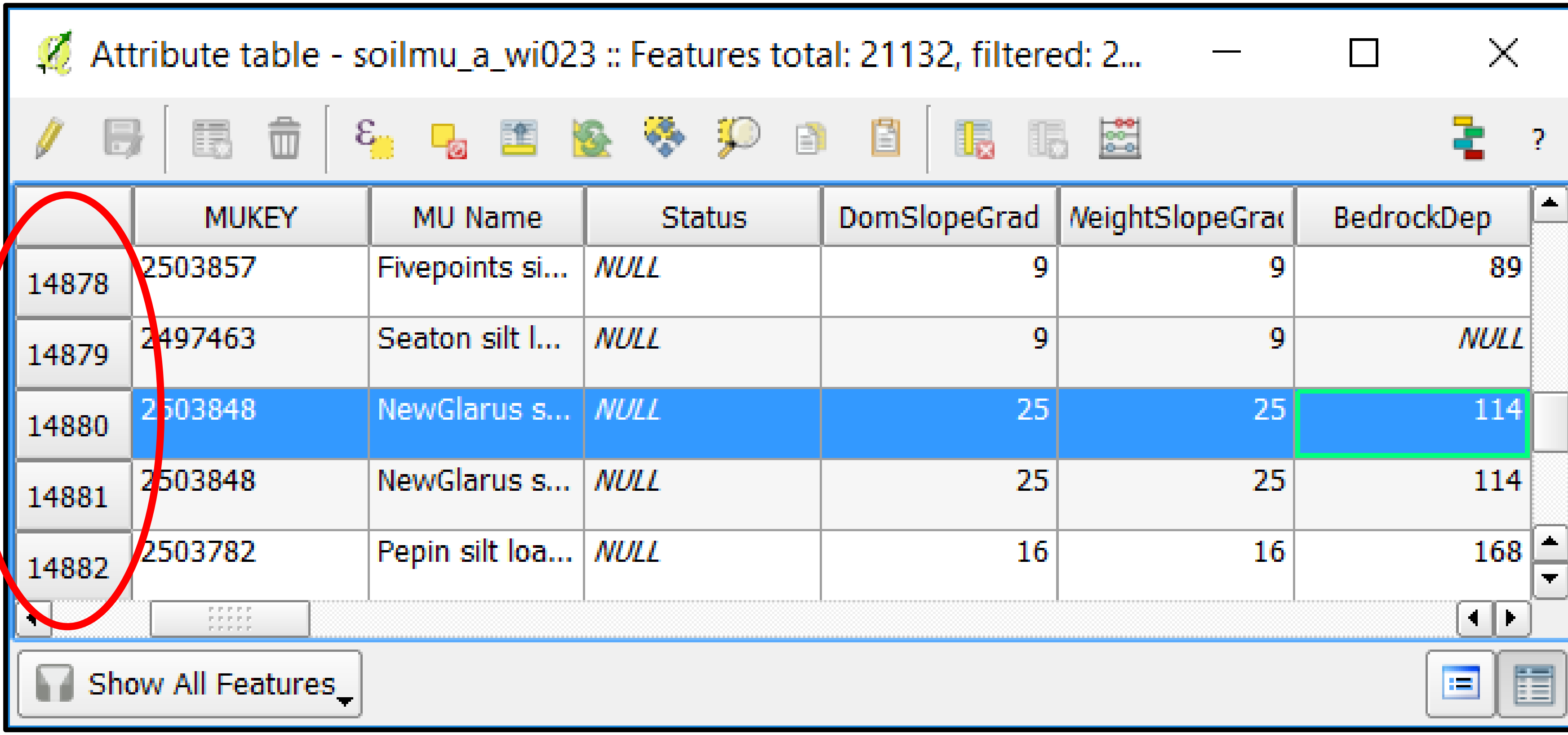
(these are attributes for each unit)



USDA Soil Survey Geographic Database (SSURGO), Crawford County

a closer look at the attribute table…

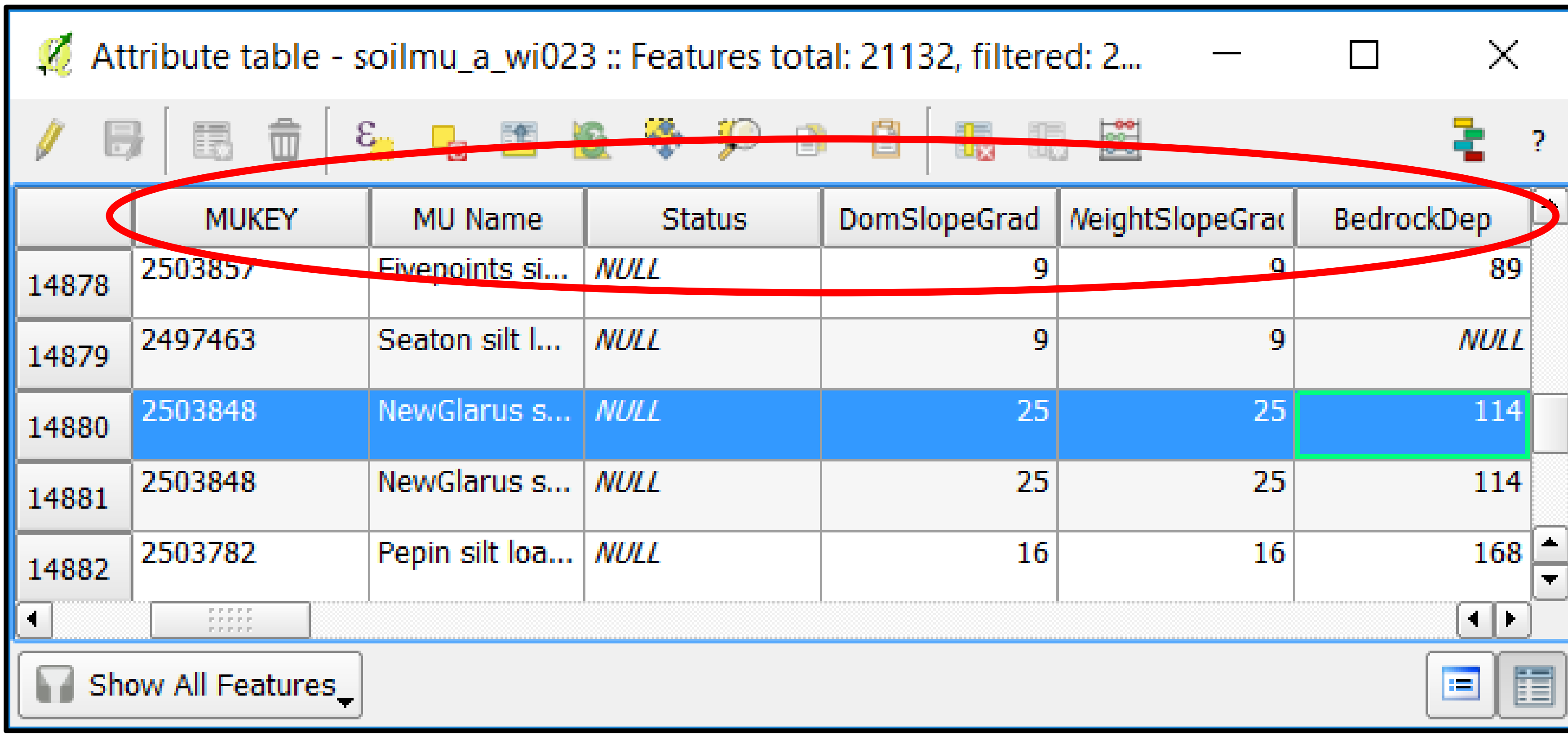
– One row per feature, each row has an “object id”



	MUKEY	MU Name	Status	DomSlopeGrad	WeightSlopeGrad	BedrockDep
14878	2503857	Fivepoints si...	NULL	9	9	89
14879	2497463	Seaton silt l...	NULL	9	9	NULL
14880	2503848	NewGlarus s...	NULL	25	25	114
14881	2503848	NewGlarus s...	NULL	25	25	114
14882	2503782	Pepin silt loa...	NULL	16	16	168

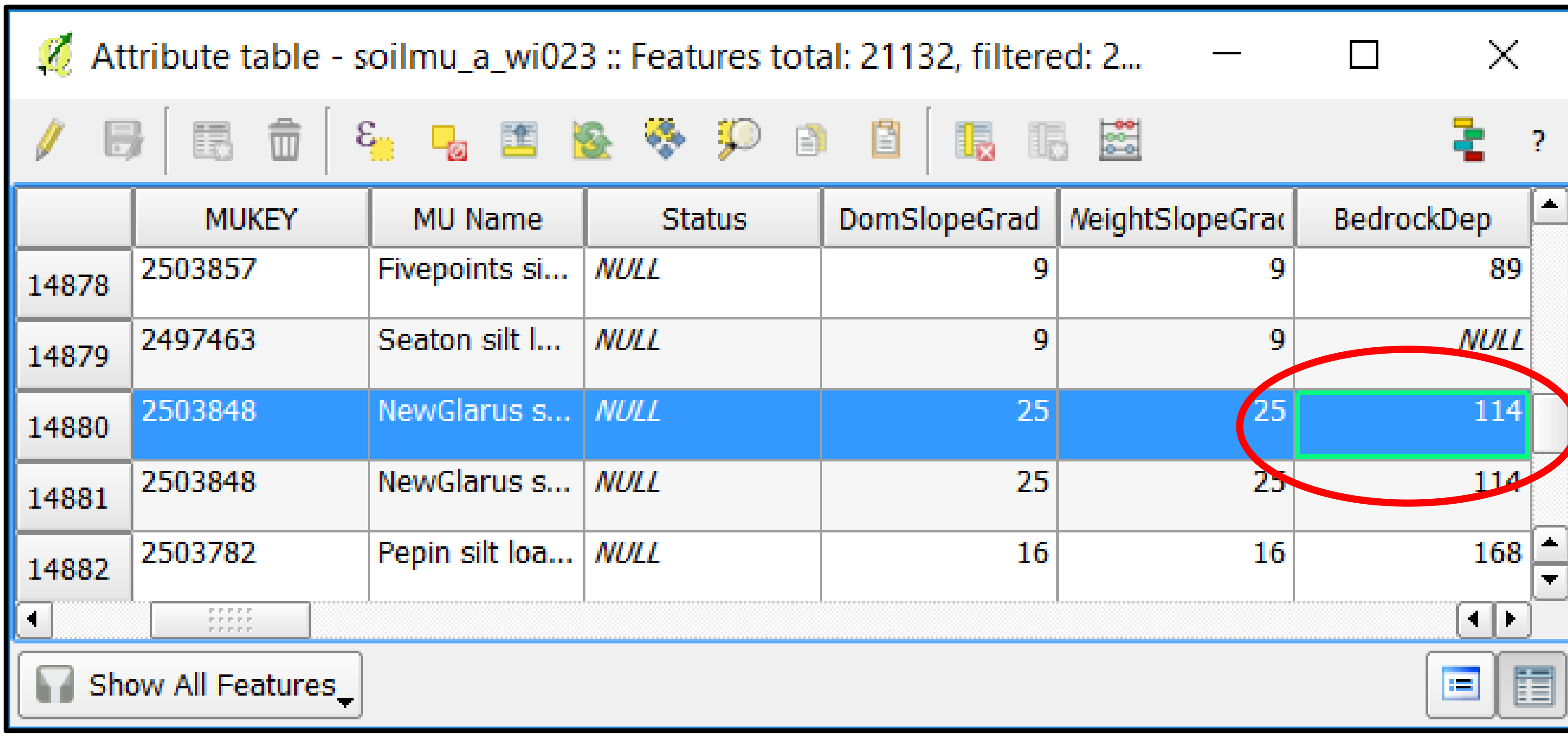
a closer look at the attribute table…

– Multiple named columns (fields)



a closer look at the attribute table…

– Here is the bedrock depth for this specific soil unit



	MUKEY	MU Name	Status	DomSlopeGrad	WeightSlopeGra	BedrockDep
14878	2503857	Fivepoints si...	NULL	9	9	89
14879	2497463	Seaton silt l...	NULL	9	9	NULL
14880	2503848	NewGlarus s...	NULL	25	25	114
14881	2503848	NewGlarus s...	NULL	25	25	114
14882	2503782	Pepin silt loa...	NULL	16	16	168

common vector formats

Delimited Text File

- Very simple text file that holds a table structure
- First line of the file holds the names of the table columns, each subsequent line holds information for one feature
- Often stored as a .csv file, or simply a .txt file

KML or KMZ file (Google Earth)

- Stores basic vector data
- Limited support for attribute table storage
- Easy use view in Google Earth, Google “My Maps”, or QGIS

common vector formats

ESRI Shapefile

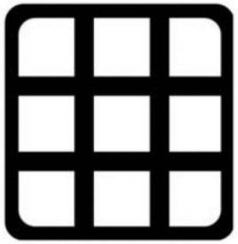
- Most ubiquitous format (and open because ESRI published the specs a while ago)
- Has some technical limitations
  - 10 character field names, blah blah…
- Basically the best thing for us to use in QGIS right now
- A single shapefile is actually a combination of a few different files:
  - .shp (this holds the geometry, and is considered the main file)
  - .shx (this is a fancy indexing thing)
  - .dbf (this is the attribute table)
  - .sbn, .cpg, .prj, .qpg, .shp.xml (these may or may not be present, and contribute indexing, projection, metadata, or other information)
  - Mapping software like QGIS aggregates all these shapefile pieces into one dataset.

common vector formats

ESRI File Geodatabase

- Currently, QGIS has read-only support for this format
- Holds vector data in “feature classes”
- NOT EXCLUSIVELY A VECTOR FORMAT: Can also hold an ESRI raster format, and annotation feature classes (independent labels)
- Can enforce “topology” rules—geometric relationships and/or constraints between feature classes

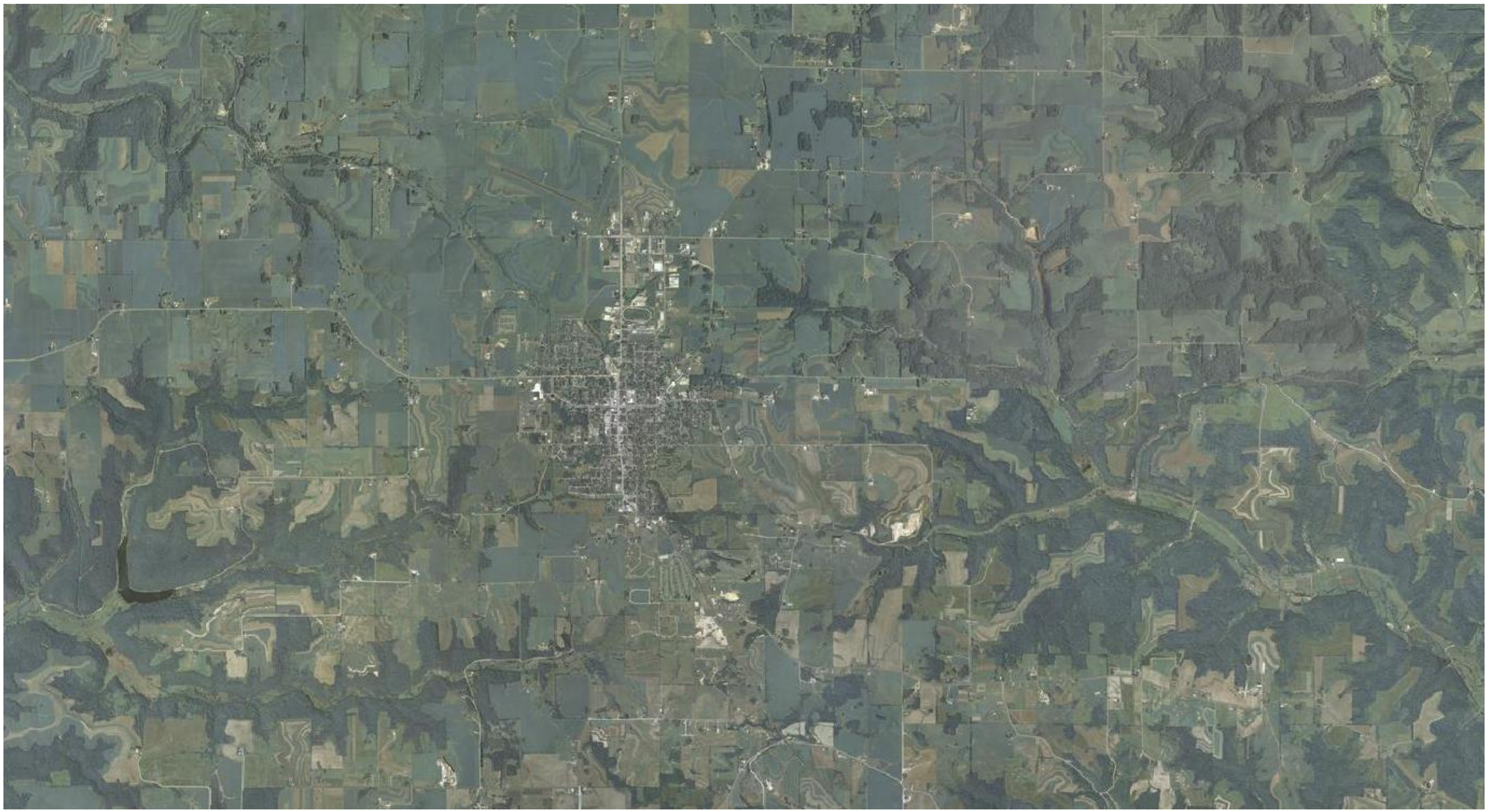
raster data



- Information stored in cells (similar to pixels)
- The “resolution” of a raster is the real-world size of each cell
  - For example, in a 30-meter resolution raster, each cell represents a square piece of the ground that is 30m x 30m
- Rasters are made up of “bands”, each band has a single value for each cell
  - Elevation rasters have 1 band, which contains the elevation value of every cell.
  - Imagery rasters generally contain 3 bands, one each for red, green and blue. When displayed properly, this gives each cell an RGB value (color) to each cell
- Use the Identify tool to view the value for any cell in a raster

raster data examples

Aerial Imagery ("orthophotos" or "ortho imagery")



MapQuest Aerial (displayed in QGIS as a Web Map Service via QuickMapServices plugin)

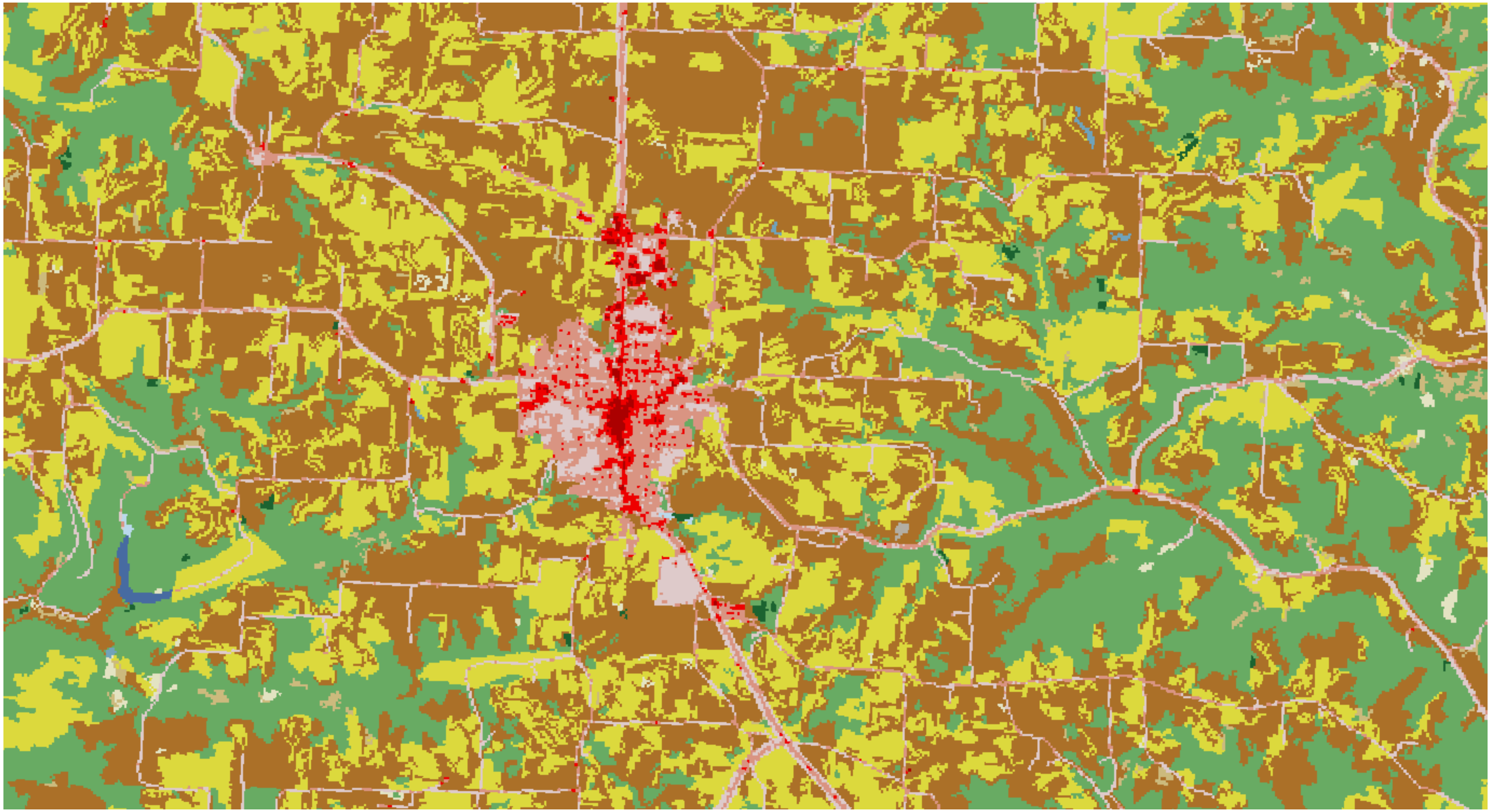
raster data examples

Elevation ("continuous data")

Vernon Co. 5ft LiDAR-derived Digital Elevation Model (found on WisconsinView)

raster data examples

Land Cover ("discrete data")



USDA National Land Cover Dataset (2011) acquired though NRCS Geospatial Data Gateway

common raster data formats

TIFF, or GeoTIFF

- "Tagged Image File Format", .tif (or .tiff)
- Embedded spatial reference or an accompanying .tfw ("world file")
- Common for discrete, or "classified", data

MrSID

- .sid, common for imagery

ESRI ArcInfo Grid

- Comes in a folder named for the raster, with files inside. In QGIS you want the file called "hdr.adf".

Image File

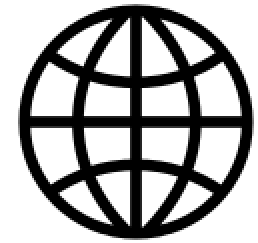
- .img, sometimes used for very large raster datasets

just for a little more perspective…

Wikipedia lists 18 different vector formats,

and 13 raster formats. I don’t know what most of them are.

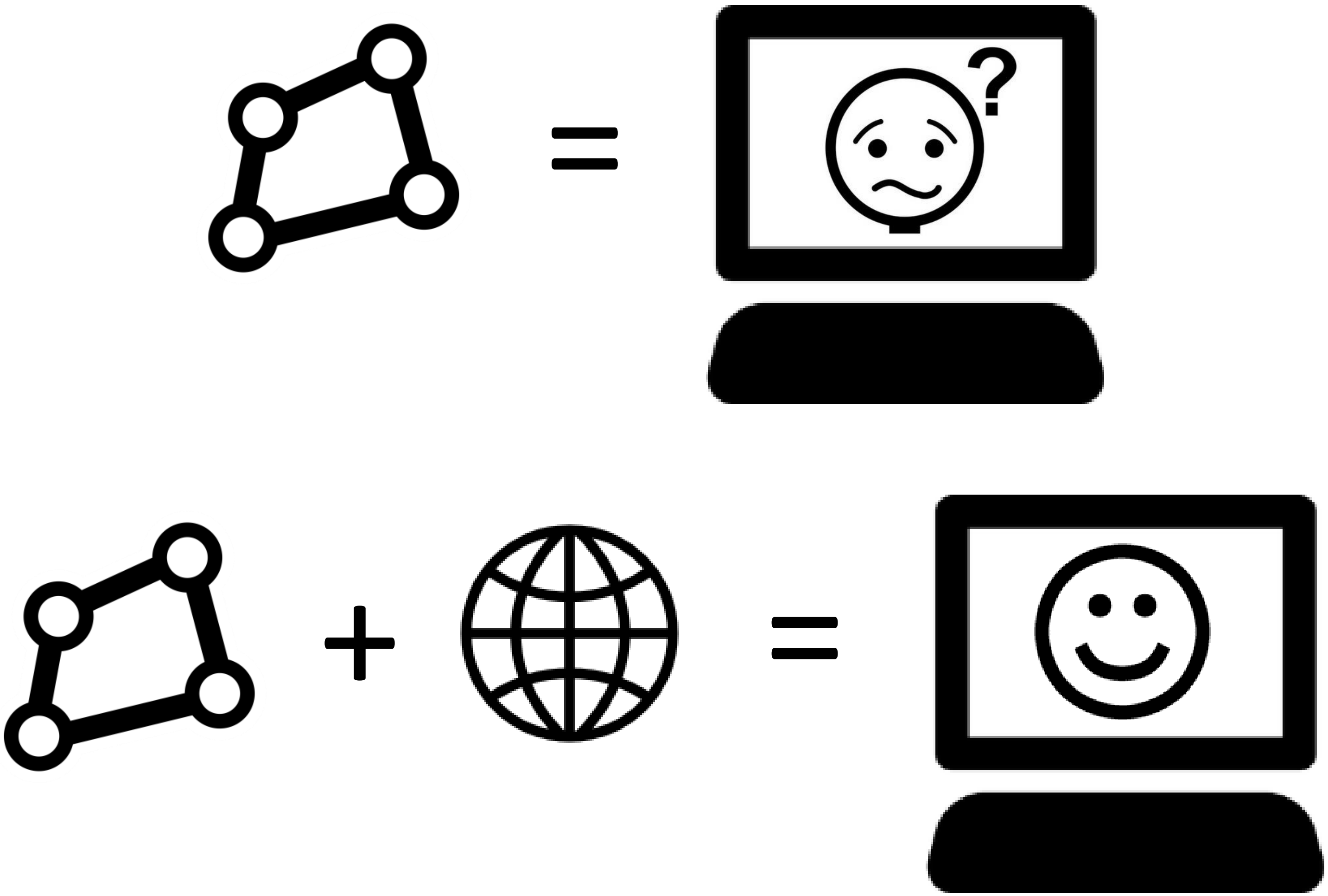
coordinate reference system (CRS)



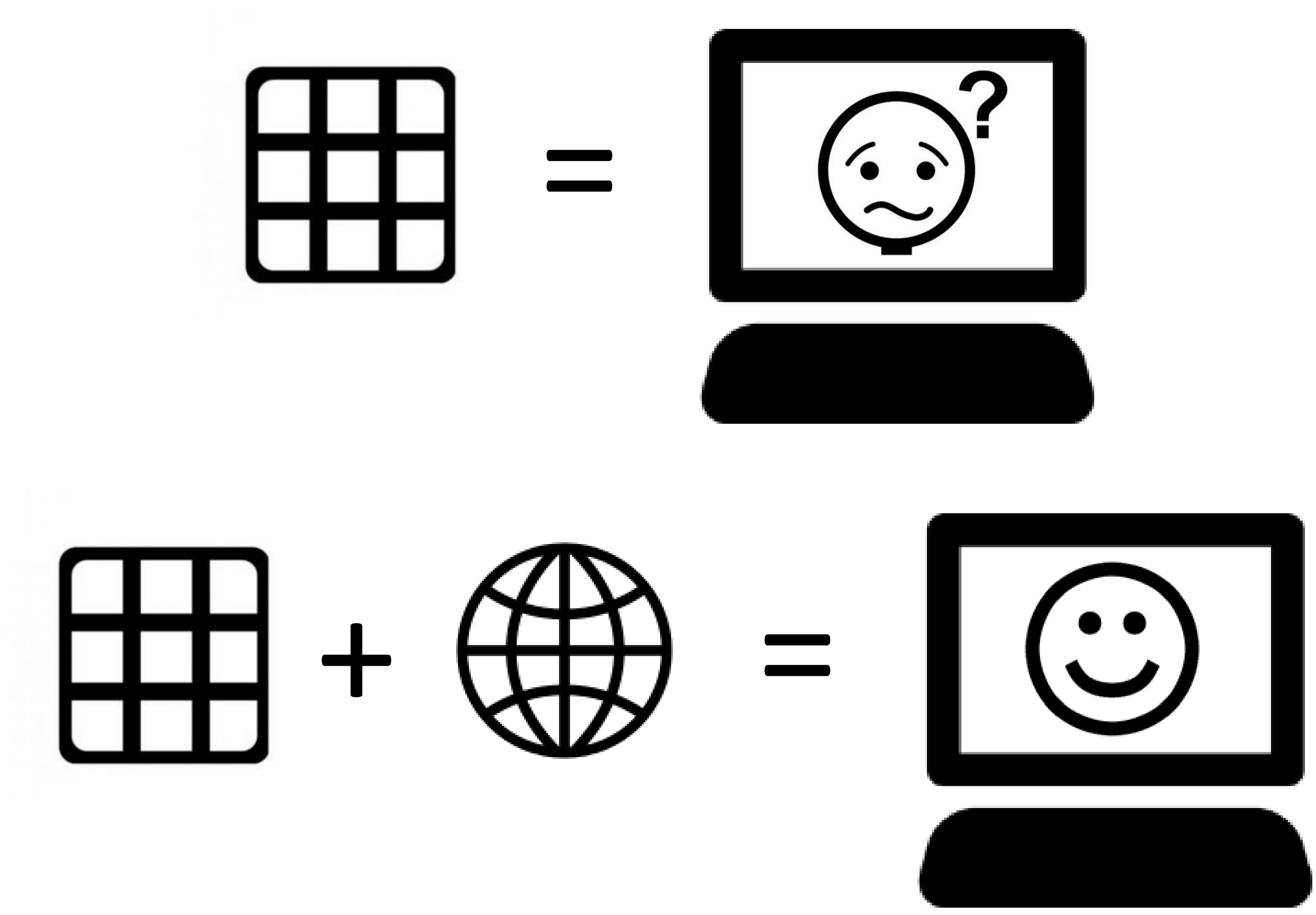
- More ways to refer to them:
  - spatial reference (SR)
  - spatial reference system (SRS)
  - projection (kind of a misnomer as you'll see)
  - We're going to stick with CRS, because that's what QGIS uses

- Where each spatial dataset is defined by its coordinates, a CRS is necessary to correctly interpret those coordinates.

A CRS tells QGIS how to interpret spatial data



A CRS tells QGIS how to interpret spatial data



coordinate reference system (CRS)

- Hundreds of different CRS… but,
- They all fall into two categories
  - **Geographic Coordinate System (GCS)**
  - **Projected Coordinate System (PCS)**
- And they are all made from the same basic elements…

elements of a CRS

First there is a datum…

…in fact, that's all you need for a **Geographic Coordinate System**.

Then a projection, an origin, and type of units are defined…

…and the result is a **Projected Coordinate System**.

elements of a CRS

First there is a datum…

datum

what is a datum?

- An abstract model of the earth's surface, based on mathematical estimations of the earth's center and the dimensions/shape of the "spheroid"
- Latitude and longitude are degree measurements around this surface
  - Latitude is a North/South measurement
  - Longitude is an East/West measurement
- Let's consider these coordinates for a moment:

  N43° 33' 24" (latitude) and W90° 53' 21" (longitude)
- They actually mark *different* points on the earth depending on which datum you use to interpret them.

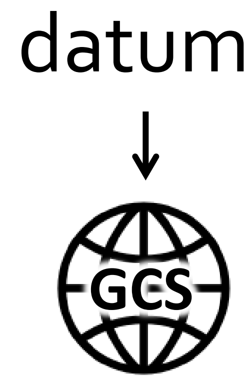
to be more specific…

- We only really care about three "datums"
  - North American Datum 1927, better known as **NAD 27**
    - Defined in 1927, you can still find it on USGS topo maps sometimes
  - North American Datum 1983, better known as **NAD 83**
    - Defined in 1983, as with NAD27, optimized to model North America as accurately as possible
  - World Geodetic Survey 1984, better known as **WGS 84**
    - Defined in 1984, used for data collection worldwide
- Other notes:
  - GPS always collect coordinates in WGS84, but you can usually set them to convert those coordinates to NAD83 (or another datum) if you want.
  - The vast majority of all *continental* US spatial data uses NAD83, maybe NAD27
  - You'll also see HARN NAD83, which is a "statewide regional upgrade of NAD83" (from the NGS NOAA website)

elements of a CRS

First there is a datum…

…in fact, that’s all you need for a **Geographic Coordinate System**.



Geographic Coordinate System recap

– A GCS interprets coordinates as latitude and longitude
– A GCS is entirely based on a single datum
– There are only three GCS of which we need to be aware:
    - NAD27
    - NAD83
    - WGS84

elements of a CRS

First there is a datum…

…in fact, that's all you need for a **Geographic Coordinate System**.

Then a projection, an origin, and type of units are defined…

datum

↓

+

projection,
origin,
units

what is a projection?

– A mathematical formula to “flatten out” the globe in order to make a 2D map

– Popular projections you may have heard of…
  - Mercator
  - Well, I guess that’s probably about it…

– Projections can get pretty wild. https://www.jasondavies.com/maps/transition/

what is an origin?

- Just like in algebra, the bottom left corner of a Cartesian grid system
- It is defined by a “central meridian” and “latitude of origin” which are included in the CRS
- It is used to focus the projection on a specific area, often called a “zone”, of the globe
  - This is because every projection can only faithfully show a certain portion of the globe (while greatly distorting the rest).

what are units in a CRS?

- Usually just feet or meters.
- On the Cartesian coordinate system created by the projection, the units are the x,y measurements.
- Example coordinates:
  - 725249,95549
  - Which means X = 725,249 and Y = 95,549
  - Or 725,249 units EAST and 95,549 units NORTH of the origin.
    - And in this CRS, units happen to be **meters**.

elements of a CRS

First there is a datum…

…in fact, that’s all you need for a **Geographic Coordinate System**.

Then a projection, an origin, and type of units are defined…

…and the result is a **Projected Coordinate System**.

Projected Coordinate System recap

- PCS interprets coordinates as an x,y position a Cartesian plane
- Units are measured in Feet or Meters (defined by the CRS itself)
- A projection is used to “flatten” the globe into a 2-dimensional plane
- Like a GCS, every PCS is based on a datum

common Projected Coordinate Systems…

– UTM Zones (Universal Transverse Mercator)

- CRS that uses UTM projection
- Use either NAD27, NAD83, or WGS84 datum
- All UTM CRS show specific zones

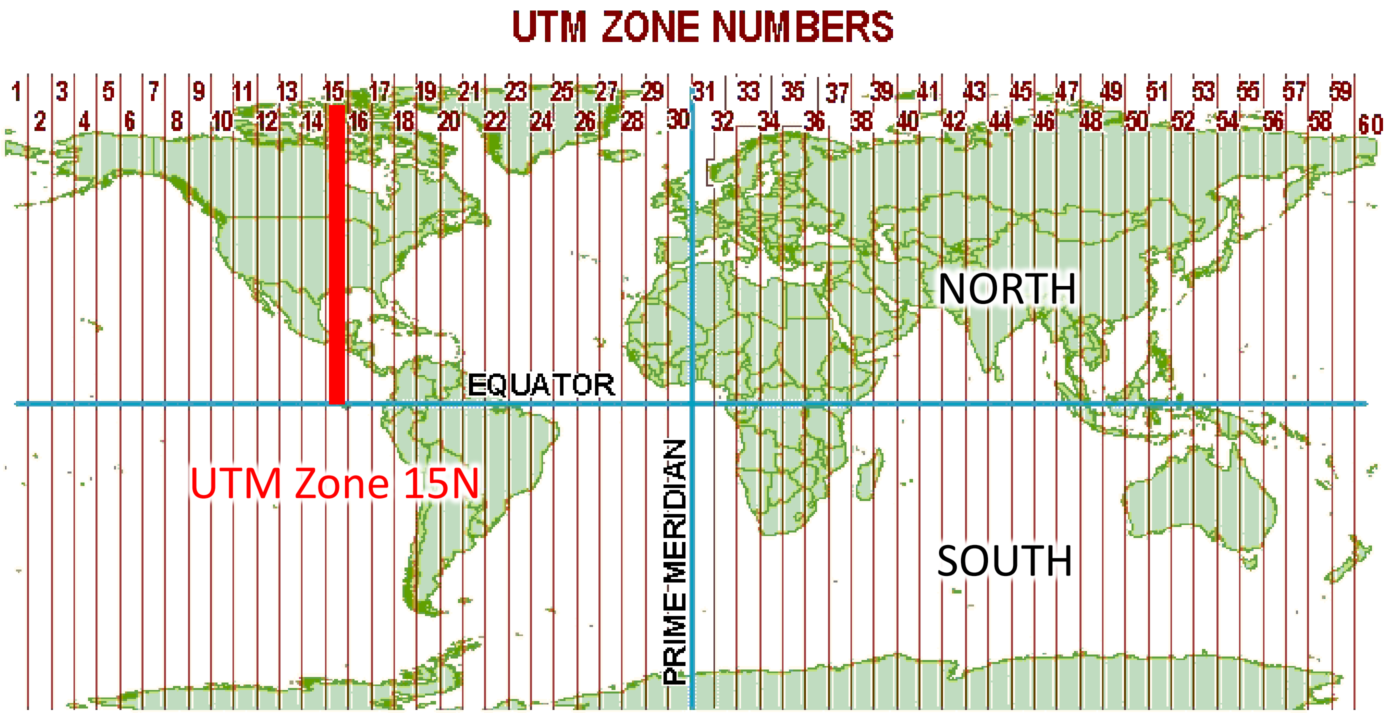


Image from National Geospatial-Intelligence Agency website

UTM Zones in the continental US

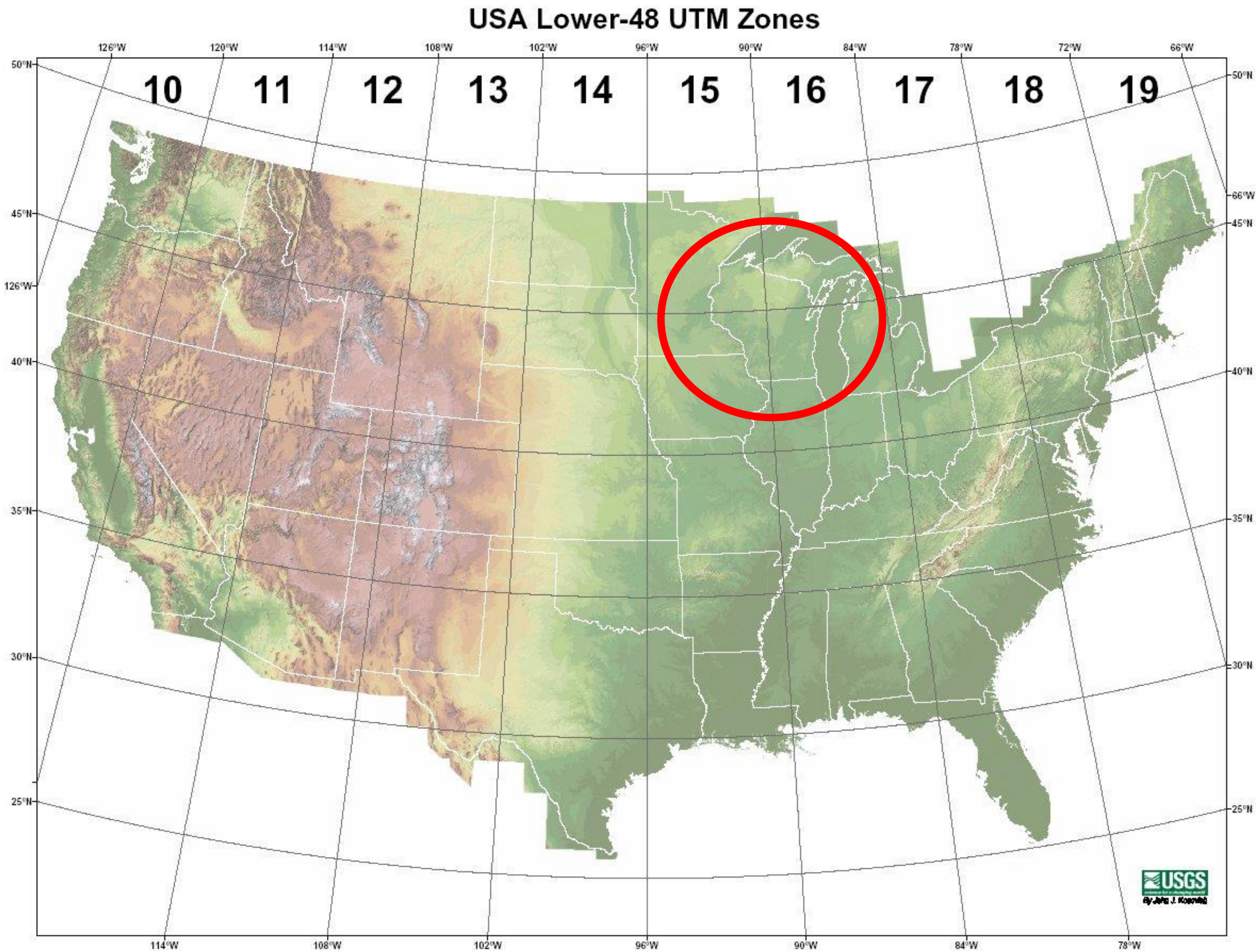


Image from National Park Service

common Projected Coordinate Systems…

– State Plane Zones
  - All of these use NAD83 datum (or NAD83 HARN)
  - Used for the best/most standard regional representations of states
  - WI has State Plane North, Central, and South

Image from National Geospatial-Intelligence Agency website

common Projected Coordinate Systems…

- Wisconsin County CRS
  - Each county in WI has its own CRS, and some county data is distributed using these.
  - You can find all of these CRS stored in projection files at this location:

  http://www.wisconsinview.org/research/projection_files/
- WGS 84 / Pseudo Mercator
  - Aka "Web Mercator", or in ArcMap "Web Mercator (Auxiliary Sphere)"
  - Used for all web maps (Google Maps, for example)

CRS Summary

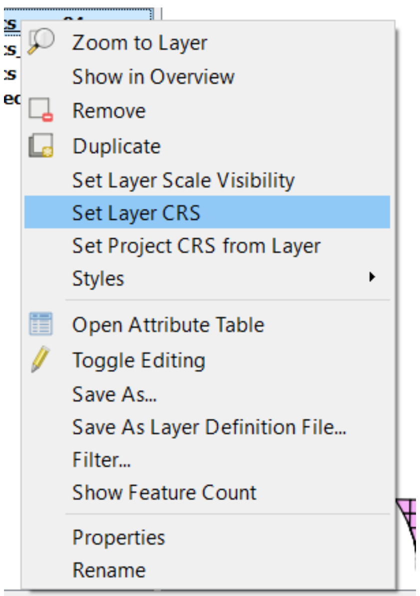
- Every spatial dataset stores coordinates
- A CRS is used to correctly interpret those coordinates
- There are two basic types of CRS:
  - Geographic Coordinate Systems
    - Interprets coordinates as latitude/longitude
  - Projected Coordinate Systems
    - Interprets coordinates as x,y positions within the Cartesian plane created by a projection and origin (as defined by a central meridian and latitude of origin)
- Final detail: All CRS have their own EPSG code (a 4-6 digit number) which you may come across

What does this mean for Quantum GIS!?

– Two different places a CRS is used:

1: At the Layer level

– Each layer must have the correct CRS defined in order for it to properly line up with all the others



» If your data has an embedded CRS or an accompanying .prj or .qpj file, you shouldn't need to do anything when adding it to your project.

» However, in many cases you'll be asked to define the CRS when you add a dataset. All you're doing is telling QGIS how to interpret your data's coordinates.

» Right-click a layer and select "Set Layer CRS" at any time

» When all layers have a correctly defined CRS, you will be happy.

What does this mean for Quantum GIS!?

– Two different places a CRS is used:

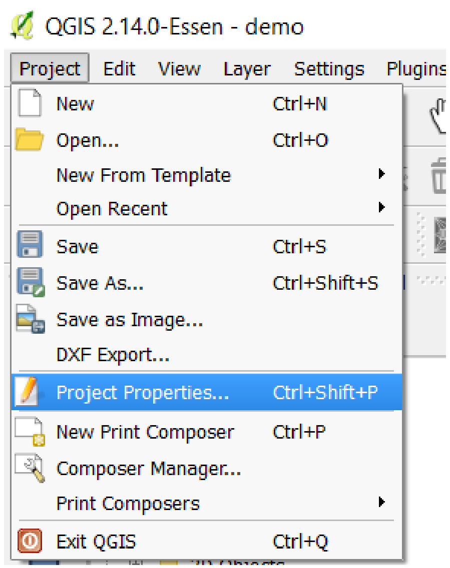
2: At the Project level

– The Project CRS is the one used for the map view

» If you are making a map of the entire US, you would want a CRS that shows the whole thing in a pleasing way (Lambert Conformal Conic, for example)

» If you are making a map of southwestern Wisconsin, you may want to use State Plane WI South

» If you are just examining data, you may simply want to use a GCS (though things will look a little squashed)



questions and break time…